

GenCore version 4.5  
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: March 1, 2001, 16:09:16 ; Search time 299.73 Seconds  
(without alignments)  
24.636 Million cell updates/sec

Title: US-09-331-631A-3\_COPY\_186\_248  
Perfect score: 353  
Sequence: 1 KRDPQOREYEDCRRCEQOE.....LINPORGSGRYEEGEEKS 63

Scoring table:  
BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 374700 seqs, 117207915 residues

Total number of hits satisfying chosen parameters: 374700

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database :  
1: SP archaea:\*  
2: SP bacteria:\*  
3: SP fungi:\*  
4: SP human:\*  
5: SP invertebrate:\*  
6: SP mammal:\*  
7: SP mhc:\*  
8: SP organelle:\*  
9: SP phage:\*  
10: SP plant:\*  
11: SP rodent:\*  
12: SP virus:\*  
13: SP vertebrate:\*  
14: SP unclassified:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

| Result No. | Score | Query Match | Length | DB ID | Description |
|------------|-------|-------------|--------|-------|-------------|
| 1          | 353   | 100.0       | 666    | 10    | Q9SP14      |
| 2          | 334   | 94.6        | 625    | 10    | Q9SP13      |
| 3          | 332   | 94.1        | 666    | 10    | Q9SP15      |
| 4          | 125.5 | 35.6        | 593    | 10    | Q9SEW4      |
| 5          | 118.5 | 33.6        | 525    | 10    | Q43358      |
| 6          | 103   | 29.2        | 637    | 10    | Q03678      |
| 7          | 89.5  | 25.4        | 810    | 10    | Q9ZM13      |
| 8          | 85.5  | 24.2        | 1483   | 5     | Q9YD46      |
| 9          | 84    | 23.8        | 236    | 10    | Q9V466      |
| 10         | 82    | 23.2        | 630    | 5     | Q9W4J3      |
| 11         | 81    | 22.9        | 242    | 10    | Q81251      |
| 12         | 81    | 22.9        | 242    | 10    | Q9SBE6      |
| 13         | 81    | 22.9        | 539    | 4     | Q9NDA2      |
| 14         | 80    | 22.7        | 238    | 10    | Q81257      |
| 15         | 80    | 22.7        | 393    | 10    | Q9ZTP0      |
| 16         | 79    | 22.4        | 238    | 10    | Q81258      |
| 17         | 79    | 22.4        | 238    | 10    | Q9SBE5      |
| 18         | 79    | 22.4        | 238    | 10    | Q9SBE6      |
| 19         | 79    | 22.4        | 242    | 10    | Q81260      |

|    |      |      |      |    |        |
|----|------|------|------|----|--------|
| 20 | 79   | 22.4 | 246  | 10 | Q81261 |
| 21 | 78.5 | 22.2 | 1655 | 5  | Q24754 |
| 22 | 78   | 22.1 | 240  | 10 | Q81252 |
| 23 | 78   | 22.1 | 240  | 10 | Q9SBE0 |
| 24 | 78   | 22.1 | 240  | 10 | Q9SBE9 |
| 25 | 78   | 22.1 | 240  | 10 | Q9SBE7 |
| 26 | 78   | 22.1 | 411  | 5  | P91419 |
| 27 | 78   | 22.1 | 540  | 10 | Q03866 |
| 28 | 77   | 21.8 | 238  | 10 | Q81255 |
| 29 | 76.5 | 21.7 | 226  | 10 | Q9S709 |
| 30 | 76.5 | 21.7 | 1128 | 5  | Q9VZ60 |
| 31 | 75.5 | 21.4 | 239  | 10 | Q9SBE1 |
| 32 | 75.5 | 21.4 | 556  | 5  | Q76940 |
| 33 | 75.5 | 21.4 | 663  | 5  | Q9W352 |
| 34 | 75   | 21.2 | 562  | 5  | Q9U345 |
| 35 | 75   | 21.2 | 568  | 5  | Q9U344 |
| 36 | 75   | 21.2 | 2123 | 5  | Q9U957 |
| 37 | 74.5 | 21.1 | 238  | 10 | Q81249 |
| 38 | 74.5 | 21.1 | 238  | 10 | Q9SBE8 |
| 39 | 74.5 | 21.1 | 407  | 10 | Q41750 |
| 40 | 74.5 | 21.1 | 582  | 10 | Q03865 |
| 41 | 74.5 | 21.1 | 686  | 5  | Q45255 |
| 42 | 74.5 | 21.1 | 1309 | 5  | Q9VDE9 |
| 43 | 74   | 21.0 | 236  | 10 | Q81250 |
| 44 | 74   | 21.0 | 236  | 10 | Q9SBE2 |
| 45 | 74   | 21.0 | 406  | 2  | Q87306 |

## ALIGNMENTS

RESULT 1  
Q9SP14 PRELIMINARY: PRT: 666 AA.  
AC Q9SP14;  
DT 01-MAY-2000 (TREMBLrel. 13, Created)  
DT 01-MAY-2000 (TREMBLrel. 13, Last sequence update)  
DT 01-OCT-2000 (TREMBLrel. 15, Last annotation update)  
DE VICILIN PRECURSOR.  
GN AMP2.  
OS Macadamia integrifolia (Macadamia nut).  
OC Eukaryota; Viridiplantae; Embryophyta; Tracheophyta; Spermatophyta;  
OC Magnoliophyta; eudicotyledons; Proteaceae; Macadamia.  
OX NCBI\_TaxID=60698;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC TISSUE=NOT KNOWN.  
RA Marcus J.P., Goulter K.C., Green J.L., Manners J.M.;  
RT "A family of antimicrobial peptides is produced by processing of a 7S  
RT globulin protein in Macadamia integrifolia.";  
RL Plant J. 0:0-0(1999).  
DR EMBL: AF161884; AAD54245.1; -  
DR HSSP: P02853; 2PHL.  
DR INTERPRO: IPR001113; -  
DR PFMW: PF00546; Seedstore\_7s; 1.  
SQ SEQUENCE 666 AA: 78243 MW: 0ECA22F8710FA7B CRC64;

Query Match 100.0% Score 353; DB 10; Length 666;  
Best Local Similarity 100.0% Pred. No. 2.2e-34;  
Matches 63; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KRDPQOREYEDCRRCEQOEPRROYOCCORRCRCEQORH\*CGDLINPORGSGRYEEGEE 60  
DB 186 KRDPQOREYEDCRRCEQOEPRROYOCCORRCRCEQORHOGDGLINPORGSGRYEEGEE 245  
QY 61 KOS 63  
DB 246 KOS 248  
RESULT 2  
Q9SP13

|    |   |              |           |                         |
|----|---|--------------|-----------|-------------------------|
| ID | 09SP13  | PRELIMINARY; | PRT;      | 625 AA.                 |
| AC | 09SP13  |              |           |                         |
| DT | 01-MAY-2000 (TReMBLrel. 13, Created)                                  |              |           |                         |
| DT | 01-MAY-2000 (TReMBLrel. 13, Last sequence update)                     |              |           |                         |
| DT | 01-OCT-2000 (TReMBLrel. 15, Last annotation update)                   |              |           |                         |
| DE | VICILIN PRECURSOR (FRAGMENT).   |              |           |                         |
| GN | AMP2.   |              |           |                         |
| OS | Macadamia integrifolia (Macadamia nut).                               |              |           |                         |
| OC | Eukaryota; Viridiplantae; Embryophyta; Tracheophyta; Spermatophyta;   |              |           |                         |
| OC | Manoliophyta; eudicotyledons; Proteaceae; Macadamia.                  |              |           |                         |
| OX | NCBI:"taxid=60698;  |              |           |                         |
| OX | [1]   |              |           |                         |
| RN | SEQUENCE FROM N.A.  |              |           |                         |
| RP | TISSUE=NUT KERNEL.  |              |           |                         |
| RC |   |              |           |                         |
| RA | Marcus J.P., Coulter K.C., Green J.L., Manners J.M.;                  |              |           |                         |
| RT | "A family of antimicrobial peptides is produced by processing of a 7S |              |           |                         |
| RT | globulin protein in Macadamia integrifolia."                          |              |           |                         |
| LT | Plant J. 0:0-0(1999).   |              |           |                         |
| DR | EMBL; AF161885; AAD54246.1; "-  |              |           |                         |
| DR | HSSP; P02853; 2PHL.   |              |           |                         |
| DR | INTERPRO; IPR001113; "-   |              |           |                         |
| DR | PFAM; PF00546; Seedstore_7s; 1.                                       |              |           |                         |
| FT | NON_TER   |              |           |                         |
| FT | 1   |              |           |                         |
| SO | SEQUENCE  | 625 AA;      | 73586 MW; | 415808A89D370296 CRC64; |

| Query Match           | 94.6% | Score 334  | DB 10 | length 625                     |
|-----------------------|-------|--|-------|--------------------------------|
| Best Local Similarity | 93.7% | Pred. No. 3,9e-32  |       |                                |
| Matches               | 59    | Conservative   | 2     | Mismatches 2; Indels 0; Gaps 0 |
| QY                    | 1     | KRPDQREYEDCRRCBDEEPPQYVQCQRRCRCEOROHMRGGDILNPPQGGSGRYEEGEE | 60    |                                |
| Db                    | 145   | KRPDQREYEDCRRCBDEEPPQYVQCQRRCRCEOROHMRGGDILNPPQGGSGRYEEGEE | 204   |                                |
| QY                    | 61    | KQS 63   |       |                                |
| Db                    | 205   | KQS 207  |       |                                |

|        |   |              |              |
|--------|---|--------------|--------------|
| RESULT | 3   |              |              |
| Q9SPL5 |   |              |              |
| ID     | Q9SPL5  | PRELIMINARY; | PRT; 666 AA. |
| AC     | Q9SPL5;   |              |              |
| DT     | 01-MAY-2000 (TREMBLrel. 13, Created)                                  |              |              |
| DT     | 01-MAY-2000 (TREMBLrel. 13, Last sequence update)                     |              |              |
| DT     | 01-OCT-2000 (TREMBLrel. 15, Last annotation update)                   |              |              |
| DE     | VICILIN PRECURSOR.  |              |              |
| GN     | AMP2.   |              |              |
| OS     | Macadamia integrifolia (Macadamia nut).                               |              |              |
| OC     | Euarystrota: Viridiplantae; Embryophyta; Tracheophyta; Spermatophyta; |              |              |
| OC     | Magnoliophyta; eudicotyledons; Proteaceae; Macadamia.                 |              |              |
| OX     | NCBI_TaxID=60698;   |              |              |
| RN     | [1]   |              |              |
| RP     | SEQUENCE FROM N.A.  |              |              |
| RC     | TISSUE=NOT KERNEL;  |              |              |
| RA     | Marcus J P., Goulter K.C., Green J.L., Manners J.M.;                  |              |              |
| RT     | giboulin protein in Macadamia integrifolia kernels."                  |              |              |
| RL     | Plant J. 0:0-0(1999).   |              |              |
| DR     | HMBL; AF161883; AAD4244.1; "  |              |              |
| DR     | HSSP; P02853; 2PHL  |              |              |
| DR     | INTERPRO: IPR001113; "  |              |              |
| PR     | PFAM; PF00546; Seedstore_7s; 1.                                       |              |              |
| SO     | SEQUENCE 666 AA; 78217 MW; C752884B2DF0224 CRC64;                     |              |              |

|                       | Query Match | Score 332   | DB 10 | Length 666                      |
|-----------------------|-------------|---|-------|---------------------------------|
| Best Local Similarity | 92.1%       | Pred. No. 7.3e-32   |       |                                 |
| Matches               | 58          | Conservative  | 4     | Mismatches 1; Indels 0; Gaps 0. |
| Qy                    | 1           | KRPDPOREYENDRRRCEDDEPRQAYOCORCRCEGDRHGRGSDLINPORGSGRYEEGEE  | 60    |                                 |
| Db                    | 186         | KRPDPOREYENDRRRCEDDEPRQAYOCORCRCEGDRHGRGSDMNPNPORGSGRYEEGEE | 245   |                                 |

|    |     |     |     |
|----|-----|-----|-----|
| QY | 61  | KQS | 63  |
|    |     | :11 |     |
| Db | 246 | EQS | 248 |

| RESULT | 4  |              |           |                         |
|--------|--|--------------|-----------|-------------------------|
| Q9SEM4 |  |              |           |                         |
| ID     | Q9SEM4   | PRELIMINARY; | PRF;      | 593 AA.                 |
| AC     | Q9SEM4;  |              |           |                         |
| DT     | 01-MAY-2000 (TREMBLrel. 13, Created)                                   |              |           |                         |
| DT     | 01-MAY-2000 (TREMBLrel. 13, Last sequence update)                      |              |           |                         |
| DT     | 01-OCT-2000 (TREMBLrel. 15, Last annotation update)                    |              |           |                         |
| DE     | VICILIN-LIKE PROTEIN PRECURSOR (FRAGMENT).                             |              |           |                         |
| OS     | Juglans regia (English walnut).  |              |           |                         |
| OC     | Eukaryota; Viridiplantae; Embryophyta; Tracheophyta; Spermatophyta;    |              |           |                         |
| OC     | Magnoliophyta; eudicotyledons; core eudicots; Rosidae; eurosids I;     |              |           |                         |
| OC     | Fagales; Juglandaceae; Juglans.  |              |           |                         |
| OX     | NCBI_TaxID=51240;  |              |           |                         |
| RN     | [1]  |              |           |                         |
| RP     | SEQUENCE FROM N.A.   |              |           |                         |
| RC     | SPRAIN-CV. SUNLAND. TISSUE-SOMATIC EMBRYO LINE;                        |              |           |                         |
| RA     | Teuber S.S., Jarvis K.C., Peterson W.R., Dandekar A.M., Ansari A.A.;   |              |           |                         |
| RT     | "Identification and cloning of a cDNA encoding a vicilin-like protein" |              |           |                         |
| RT     | Jug r 2, from English walnut kernel (Juglans regia) a major food       |              |           |                         |
| RT     | allergen."   |              |           |                         |
| RL     | Submitted (MAY-1998) to the EMBL/Genbank/DBU databases.                |              |           |                         |
| DR     | EMBL; AF066605; AAF18269.1; -  |              |           |                         |
| DR     | HSSP; P02853; ZPHL.  |              |           |                         |
| DR     | INTERPRO: IPR001113; -   |              |           |                         |
| DR     | PFAM; PF00546; Seedstore_7s; 1.  |              |           |                         |
| FT     | NON_TER  |              |           |                         |
| FT     | 1  |              |           |                         |
| QO     | SEQUENCE   | 593 AA;      | 69990 MW; | 9BA127E19B18C0D8 CRC64; |

|             |              |  |            |         |        |    |        |      |
|-------------|--------------|--|------------|---------|--------|----|--------|------|
| Query Match | Similarity   | 35.6%  | Score      | 125.5   | DB     | 10 | Length | 593; |
| Best Local  | Similarity   | 45.9%  | Pred. No.  | 3.3e-07 |        |    |        |      |
| Matches     | Conservative | 28;  | Mismatches | 16;     | Indels | 5; | Gaps   | 4;   |
| OY          | 1            | KRDPOQREYDCRRRC--EQDEPRQOYLQCQRRC-RQQQHQRH-GGDLINPQGSGGRYE | 56         |         |        |    |        |      |
| Dd          | 116          | RDDPD--QYHRQCRKQIQIOESPERORRQDQRCEROYKEQGSFRERGPSPRESRGREE | 174        |         |        |    |        |      |
| OY          | 57           | E 57   |            |         |        |    |        |      |
| Dd          | 175          | E 175  |            |         |        |    |        |      |

| Result | 5  |              |      |         |
|--------|--|--------------|------|---------|
| Q43358 |  |              |      |         |
| ID     | Q43358   | PRELIMINARY; | PRT; | 525 AA. |
| AC     | Q43358;  |              |      |         |
| DT     | 01-NOV-1996 (TrEMBLrel. 01, Created)                                   |              |      |         |
| DT     | 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)                      |              |      |         |
| DT     | 01-OCT-2000 (TrEMBLrel. 15, Last annotation update)                    |              |      |         |
| DE     | VICILIN PRECURSOR.   |              |      |         |
| GN     | CSV.   |              |      |         |
| OS     | Theobroma cacao (Cacao).   |              |      |         |
| OC     | Eukaryota; Viridiplantae; Embryophyta; Tracheophyta; Spermatophyta;    |              |      |         |
| OC     | Magnoliophyta; eudicots; core eudicots; Rosidae; eurosids II;          |              |      |         |
| OC     | Malvales; Malvaceae; Theobroma.  |              |      |         |
| OX     | NCBI_TaxID=3641;   |              |      |         |
| RN     | 11   |              |      |         |
| RP     | SEQUENCE FROM N.A.   |              |      |         |
| RC     | TISSUE=LEAVES;   |              |      |         |
| RX     | MEDLINE=92288309; Pubmed=1600151;                                      |              |      |         |
| RA     | McHenry L., Fritz P.J.;  |              |      |         |
| RT     | "Comparison of the structure and nucleotide sequences of vicilin genes |              |      |         |
| RT     | of cocoa and cotton raise questions about vicilin evolution.";         |              |      |         |
| RL     | Plant Mol. Biol. 18:1173-1176(1992).                                   |              |      |         |
| DR     | EMBL; X62625; CAA44493.1; -  |              |      |         |
| DR     | EMBL; X62626; CAA44494.1; -  |              |      |         |
| DR     | HSSP; P02853; ZPHL.  |              |      |         |

|              |         |   |              |                       |            |           |             |     |
|--------------|---------|---|--------------|-----------------------|------------|-----------|-------------|-----|
|              |         |   |              | 29.2%                 | Score 103; | DB 10;    | Length 637; |     |
|              |         |   |              | Best local Similarity | 31.9%;     | Pred. No. | 0.00017;    |     |
|              | Matches | 22;   | Conservative | 11;                   | Mismatches | 20;       | Indels      | 16; |
|              |         |   |              |                       |            | Gaps      |             | 2   |
| QY           | 10      | EDCRRRCCEQOEPRQY-OCORRCREOPRONGSDLNPORG-----GSG 53                      |              |                       |            |           |             |     |
|              |         | :   :     :   :   :   :   :   :   :   :   :   :   :   :   :   :   :   : |              |                       |            |           |             |     |
| Db           | 42      | OCCVCNCRGERBRYSHARCVDCCRDDQQDQHNRHQEEEGQGRGWHEGEEREEHNGRG 101           |              |                       |            |           |             |     |
| QY           | 54      | KYELEGEEKO 62   |              |                       |            |           |             |     |
|              |         | :     : :   |              |                       |            |           |             |     |
| Db           | 102     | RHGEGERE 110  |              |                       |            |           |             |     |
| RESULT       | 7       |   |              |                       |            |           |             |     |
| ID           | Q9ZMI3  |   |              |                       |            |           |             |     |
| PRELIMINARY: |         |   |              |                       |            |           |             |     |
| PRT:         |         |   |              |                       |            |           |             |     |
| 810 AA.      |         |   |              |                       |            |           |             |     |

RP SEQUENCE FROM N.A.  
RC STRAIN-BERKELEY;  
RX MEDLINE-20196006; PubMed=10731132;  
RA Adams M.D., Celniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,  
RA Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galle R.F.,  
RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,  
RA Sultion G.G., Mortman J.R., Yandell M.D., Zhang Q., Chen L.X.,  
RA Brandon R.C., Rogers J.H.C., Blatej R.G., Champe M., Pfeiffer B.D.,  
RA Wan R.H., Doyle C., Baxter E.G., Hertz G., Nelson C.R., Miklos G.L.G.,  
RA Aillet J.F., Agayanyi A., An H.-J., Andrews-Pfannkoch C., Baldwin D.,  
RA Baillif R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,  
RA Beeson K.Y., Benos P.V., Betman B.P., Bhandari D., Bolshakov S.,  
RA Borokova D., Botchan M.R., Bouck J., Brokstein P., Brotler P.,  
RA Burits K.C., Busam D.A., Butler H., Cardaud E., Center A., Chandra I.,  
RA Cherry J.M., Chewley S., Dahlke C., Davenport L.B., Davies P.,  
RA de Pablos B., Delcher A., Deng Z., May's A.D., Dew I., Dietz S.M.,  
RA Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,  
RA Durbin K.J., Evangelista C.C., Ferraz C., Ferreira S., Fleischmann W.,  
RA Foster C., Gabrielian A.E., Gang N.S., Gelbart W.M., Glasser K.,  
RA Glodok A., Gong F., Gottrell J.H., Gu Z., Guan P., Harris M.,



| RESULT | ID  | PRELIMINARY; | PRT; | 242 AA. |
|--------|---|--------------|------|---------|
| 11     | 081251  |              |      |         |
| AC     | 081251;   |              |      |         |
| DT     | 01-NOV-1998 (TREMBLrel. 08, Created)                                |              |      |         |
| DT     | 01-NOV-1998 (TREMBLrel. 08, Last sequence update)                   |              |      |         |
| DT     | 01-OCT-2000 (TREMBLrel. 15, Last annotation update)                 |              |      |         |
| DE     | GLIOBULIN-1 (FRAGMENT).   |              |      |         |
| OS     | Zea mays subsp. mays (maize).                                       |              |      |         |
| OC     | Eukaryota; Viridiplantae; Embryophyta; Tracheophyta; Spermatophyta; |              |      |         |
| OC     | Magnoliophyta; Liliopsida; Poales; Poaceae; Zea.                    |              |      |         |
| OX     | NCBI_TaxID=4578;  |              |      |         |
| RN     | [1]   |              |      |         |
| RP     | SEQUENCE FROM N.A.  |              |      |         |
| RA     | Hilton H., Gaut B.S.;   |              |      |         |
| RT     | "Speciation and domestication in maize and its wild relatives:      |              |      |         |
| RT     | Genetics 0:0-0(1998)."  |              |      |         |
| RL     | EMBL; AF064214; AAC31457.1; -.                                      |              |      |         |
| DR     | HSSP; P50477; ICAU.   |              |      |         |
| DR     | MENDEL; 31893; Zeamaa;1188;31893.                                   |              |      |         |
| DR     | INTERPRO; IPR000901; -.   |              |      |         |
| DR     | INTERPRO; IPR001113; -.   |              |      |         |
| DR     | PFAM; PF00546; Seedsstore_7s; 1.                                    |              |      |         |
| DR     | PROSITE; PS00867; CPSSSE_2; UNKNOWN_1.                              |              |      |         |
| FT     | NON_TER 242   |              |      |         |
| SO     | SEQUENCE 242 AA; 27860 MW; 913912DE924B0ADB CRC64;                  |              |      |         |

|                       |                 |                 |            |             |
|-----------------------|-----------------|-----------------|------------|-------------|
| Query Match           | 22.9%;          | Score 81;       | DB 10;     | Length 242; |
| Best Local Similarity | 35.9%;          | Pred. NO. 0.03; |            |             |
| Matches 23;           | Conservative 7; | Mismatches 18;  | Indels 16; | Gaps 3;     |

Db 40 CVRCEERPMHQPCLEQCEKREKREKROEGRHE--TDRSRGEGSSIEDEREREREDEK 96

|                   |   |              |                                   |
|-------------------|---|--------------|-----------------------------------|
| RESULT            | 12  |              |                                   |
| 09SBE6            |   |              |                                   |
| ID                | 09SBE6  | PRELIMINARY; | PRT; 242 AA.                      |
| AC                | 09SBE6;   |              |                                   |
| DT                | 01-MAY-2000 (Tremblrel, 13, Created)                                |              |                                   |
| DT                | 01-MAY-2000 (Tremblrel, 13, last sequence update)                   |              |                                   |
| DT                | 01-OCT-2000 (Tremblrel, 15, last annotation update)                 |              |                                   |
| DE                | GLOBALIN-1 (FRAGMENT).  |              |                                   |
| OS                | Zea mays subsp. parviglumis.  |              |                                   |
| OC                | Eukaryota; Viridiplantae; Embryophyta; Tracheophyta; Spermatophyta; |              |                                   |
| OC                | Magnoliophyta; Liliopsida; Poales; Poaceae; Zea.                    |              |                                   |
| NCBI_taxid=76912; | [1]   |              |                                   |
| RP                | SEQUENCE FROM N.A.  |              |                                   |
| RA                | Hilton H., Gaut B. S.;  |              |                                   |
| RT                | "speciation and domestication in maize and its wild relatives:      |              |                                   |
| RT                | evidence from the Globulin-1 gene."                                 |              |                                   |
| TL                | Genetics 0:0-0(1998).   |              |                                   |
| DR                | EMBL; AF064425; AAC31468.1;   |              |                                   |
| DR                | HSSP; P50477; ICAU.   |              |                                   |
| DR                | INTERPRO; IPR001113.  |              |                                   |
| DR                | PRAM; PR00546; Seedstore-7s; 1.                                     |              |                                   |
| FT                | NON_TER   | 242          | 242                               |
| Q                 | SEQUENCE  | 242 AA;      | 27856 MW; 866DB2C9E45C6BAD CRC64; |

|                       |        |                 |        |                |
|-----------------------|--------|-----------------|--------|----------------|
| Query Match           | 22.9%; | Score 81;       | DB 10; | Length 242;    |
| Best Local Similarity | 35.9%; | Pred. No. 0.03; |        |                |
| Matches               | 23;    | Conservative    | 7;     | Mismatches 18; |
|                       |        |                 |        | Indels 16;     |
|                       |        |                 |        | Gaps 3;        |

DQ  
DB

97 EKKO 100

59 EKKO 62

IIII

QY

## 13

DT 01-OCT-2000 (T-EMBLrel. 15, Created)  
DT 01-OCT-2000 (T-EMBLrel. 15, last sequence update)  
DT 01-OCT-2000 (T-EMBLrel. 15, last annotation update)  
DE D138084.1 (ANDROGEN RECEPTOR (DIHYDROTESTOSTERONE RECEPTOR))  
DE (FRAGMENT).  
GN AR.  
OS Homo sapiens (Human).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
OX NCBI\_TaxID=9606;  
RN 11  
RP SEQUENCE FROM N.A.  
RA Chapman J.;  
RL Submitted (APR-2000) to the EMBL/GenBank/DBJ databases.  
DR EMBL: AL049564; CAB87955.1; -.  
KW Receptor.  
FT 539 539  
EQ NONTER.  
SQ SEQUENCE 539 AA; 55444 MW; AB493953B89D869F CRC64;

|             |        |           |       |             |
|-------------|--------|-----------|-------|-------------|
| Query Match | 22.98; | Score 81; | DB 4; | Length 539; |
|-------------|--------|-----------|-------|-------------|

|    | Matches | 19,         | Conservative | 15;      | Mismatches | 24;      | Indels  | 0;     | Gaps        | 0;  |
|----|---------|-------------|--------------|----------|------------|----------|---------|--------|-------------|-----|
| QY | 5       | QQREVEDRRRC | EQQEP        | QQOYOCQR | KCRHQDQH   | RGGDLINP | OKSGSRV | EEGEEK | 62          |     |
|    |         | :           | :            | :        | :          | :        | :   :   | :      | :           |     |
| Dd | 58      | QQQQQQQQQQ  | QQQQQQQQ     | QQQQQQQQ | EQTSPP     | QQQQQQQ  | GEDSDSP | QAHRRG | PPTGLVLDEEQ | 115 |

| RESULT   | ID        | PRELIMINARY; | PRT; | 238 AA. |
|--|-----------|--------------|------|---------|
| 081257   | ID 081257 |              |      |         |
| AC 081257;   |           |              |      |         |
| DT 01-NOV-1998 (TREMBLrel. 08, Created)                                |           |              |      |         |
| DT 01-NOV-1998 (TREMBLrel. 08, last sequence update)                   |           |              |      |         |
| DT 01-OCT-2000 (TREMBLrel. 15, last annotation update)                 |           |              |      |         |
| DE GLOBULIN-1 (FRAGMENT),  |           |              |      |         |
| OS Zea luxurians (Teosinte).   |           |              |      |         |
| OC Eukaryota; Viridiplantae; Embryophyta; Tracheophyta; Spermatophyta; |           |              |      |         |
| OC Magnoliophyta; Liliopsida; Poales; Poaceae; Zea.                    |           |              |      |         |
| NCBI_TaxID=15945;  |           |              |      |         |
| [1]  |           |              |      |         |
| SEQUENCE FROM N.A.   |           |              |      |         |
| RA Hilton H., Gaut B.S.;   |           |              |      |         |
| RT "Speciation and domestication in maize and its wild relatives:      |           |              |      |         |
| RT evidence from the Globulin-1 gene.";                                |           |              |      |         |
| RL GenBank: 0-0-0(1998).   |           |              |      |         |
| DR EMBL: AF064328; AAC31471.1; -.                                      |           |              |      |         |
| DR HSSP: P50477; 1CAU.   |           |              |      |         |
| DR MENDEL: 31899; Zea1u;1186;31899.                                    |           |              |      |         |
| DR INTERPRO: IPR000901; -.   |           |              |      |         |
| DR INTERPRO: IPR001113; -.   |           |              |      |         |
| DR PFAM: PF00546; Seedstore_7s; 1.                                     |           |              |      |         |
| DR PROSITE: PS00867; CPAS5E_2; UNKNOWN_1.                              |           |              |      |         |
| FT NON_TER 238   |           |              |      |         |
| EQ SEQUENCE 238 AA: 27233 MW; FD31A8066739E68 CRC64;                   |           |              |      |         |

|                       |       |           |        |             |
|-----------------------|-------|-----------|--------|-------------|
| Query Match           | 22.7% | Score 80; | DB 10; | Length 238; |
| Best Local Similarity | 35.5% | Pred. No. | 0.039; |             |

Matches 22; Conservative 6; Mismatches 18; Indels 16; Gaps 3;

QY 12 CRRCEQOEPQOYOCORCR-----OQOHGRCGLINPQRCGSG-----RYEGEE 60  
| | | | : | : | | | : | | | : | | | : | | |  
DB 40 CVRCEDRPWQHQRCLQCRFEERERKQDRSRHQAAD-----RSQGSSEDERQKEE 94

QY 61 KQ 62  
| |  
DB 95 KQ 96

## RESULT 15

Q9ZTP0 PRELIMINARY; PRT; 393 AA.

ID Q9ZTP0;  
AC Q9ZTP0;  
DT 01-MAY-1999 (TrEMBLrel. 10, Created)  
DT 01-MAY-1999 (TrEMBLrel. 10, Last sequence update)  
DT 01-MAY-1999 (TrEMBLrel. 10, Last annotation update)  
DE HYPOTHETICAL 45.3 KDA PROTEIN.  
GN OSE705.  
OS Oryza sativa (Rice).  
OC Eukaryota; Viridiplantae; Embryophyta; Tracheophyta; Spermatophyta;  
OC Magnoliophyta; Liliopsida; Poales; Poaceae; Oryza.  
OX NCBI\_TaxID=4530;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=LONELLO;  
RA Chen P.W., Chen L.J.;  
RL Submitted (FEB-1998) to the EMBL/Genbank/DBJ databases.  
DR EMBL: AF049348; AAD02494.1; -.  
KW Hypothetical protein.  
SQ SEQUENCE 393 AA; 45258 MW; DBD01934BA2F9E95 CRC64;

Query Match 22.7%; Score 80; DB 10; Length 393;

Best Local Similarity 31.8%; Pred. No. 0.062; Indels 26; Gaps 4;  
Matches 21; Conservative 7; Mismatches 12;

QY 6 QREYED-----CRRCE--QOEPQOYOCORC---REOQOHGRCGLIN 46  
| | | | : | : | | | : | | | : | | | : | | |  
DB 67 ORQOEDDDDDMTGGGCKRCRCRCRGYRDEPWKQECMRQCEWRRHQHNGSG----- 121

QY 47 PQRGGS 52  
| | |  
DB 122 --HGS 125

Search completed: March 1, 2001, 16:09:17  
Job time: 1561 sec